

Visuospatial Bootstrapping: Long Term Memory Representations are Necessary for Implicit
Binding of Verbal and Visuospatial Working Memory

Stephen Darling^a, Richard J. Allen^b, Jelena Havelka^b, Aileen Campbell^a & Emma Rattray^a

^a - Queen Margaret University, Edinburgh

^b - Institute for Psychological Sciences, University of Leeds

Word Count: 3966

Correspondence concerning this article should be addressed to:

Dr Stephen Darling

Division of Psychology and Sociology

Queen Margaret University

Edinburgh

EH21 6UU

E: sdarling@qmu.ac.uk

Tel: 0131 474 0000

Abstract

It has recently been shown that presenting additional visuospatial information alongside to-be-remembered numbers in a digit span task enhances participants' memory for those items. However, the mechanisms behind this 'visuospatial bootstrapping' effect remained unspecified. In this paper we report evidence that this effect involves an integration of information from verbal and visuospatial temporary memory with long-term-memory representations, and that the existence of a relevant LTM representation is necessary for bootstrapping to occur.

Keywords: Working memory; Visuospatial working memory; Phonological working memory; Episodic buffer.

Visuospatial Bootstrapping: Long Term Memory Representations are Necessary for Implicit Binding of Verbal and Visuospatial Working Memory

Early theoretical frameworks of Working Memory (e.g. Baddeley & Hitch, 1974) assert that visuospatial working memory (VSWM) and verbal working memory (VWM) represent discrete subsystems. There is a range of evidence supporting the segregation of verbal and visuospatial temporary memory systems (e.g. Smith, Jonides, & Koeppel, 1996; Baddeley, Lewis, & Vallar, 1984; Quinn & McConnell, 1996; Cocchini, Logie, Della Sala, MacPherson & Baddeley, 2002). Nonetheless, there is also evidence that under certain circumstances verbal and visuospatial working memory do not operate entirely independently: for example, Morey and Cowan (2004, 2005; see also Mate, Allen, & Baques, in press) reported a number of studies where verbal articulations impaired memory on a VSWM task. These findings point to the involvement of domain-general processes in ostensibly domain-specific Working Memory (WM) tasks, and hence suggest a limit on the independence of WM subsystems.

In order to account for findings like this, original models of the relationship between working memory subsystems have been modified to incorporate a new component, the 'Episodic Buffer' (EB: Baddeley, 2000; Baddeley, Allen & Hitch, 2011). This represents a limited capacity store that is recruited when information from different sources, including the WM subsystems and LTM, has to be bound together and temporarily retained.

An alternative and influential theoretical perspective is offered by Cowan (2005, 1998, 1988), in which working memory processes are conceived of as interactive activations of LTM and STM that fall within a focus of attention. Within such a model, there is no specific need to segregate component memory processes such as a visuo-spatial sketch pad or phonological loop, and thus no need for an EB: Cowan, Saults and Morey (2006) argue that when task demands require retention of combinations of verbal and spatial material (such as

‘Ann lives here’), this is greatly facilitated (in adults) by parallel matching based on serial order. However, they also identified a residual capacity for representations combining verbal and spatial materials, consistent with Cowan’s notion of a focus of attention, and also with Baddeley’s episodic buffer.

Baddeley’s initial conception of the Episodic Buffer (Baddeley, 2000) was a component that required executive resources to function. However, research has not supported that position, with evidence of automatic binding in both verbal and visuospatial memory (Allen, Hitch & Baddeley, 2009; Baddeley, Hitch & Allen, 2009). Consequently the latest description of the working memory model (Baddeley et al., 2011) characterizes many forms of binding as developing within the EB or feeding into it from other subcomponents of WM, in both cases without the intervention of executive attentional processes. Instead, basic, rule-governed and automatic filtering mechanisms are proposed which function to select which features are held as bound representations (Ueno, Mate, Allen, Hitch & Baddeley, 2011), with long-term memory support capable of impacting on WM representations independently of attentional control (Baddeley et al., 2011). This specific aspect of the EB – its lack of requirement for attention – is a potential point of distinction between Baddeley’s approach and Cowan’s: the focus of attention model implies a more important role for explicit attention in feature binding of this type, and the question of whether attention is critical in binding remains disputed due to reports of contradictory findings, such as the observation that binding between individual letters and locations was reduced by a tone memory load (Elsley & Parmentier, 2009).

A surprisingly limited amount of research has focused on binding of verbal and spatial material, despite the fact that the distinction of VSWM and VWM is one of the key features of the multicomponent working memory model. Nonetheless, there is evidence that verbal and spatial information can be bound together (Bao, Li & Zhang, 2007; Prabhakaran,

Narayanan, Zhao & Gabrieli, 2000). Binding has also been observed between colour *words* and shapes (Allen, et al., 2009). Recently, our group has reported interactions between verbal and visuospatial working memory in task that does not *explicitly* ask participants to bind information, and that therefore demonstrates the presence of implicit cross-modality feature binding. We refer to an effect termed ‘visuospatial bootstrapping’, which was observed on a visually presented verbal digit recall task (Darling & Havelka, 2010). Participants were shown sequences of digits for immediate recall; when these were presented by highlighting a sequence of digits in a spatial array that was arranged as a traditional telephone keypad, digits were better remembered than if they were presented as single digits in the centre of the screen. This was assumed to reflect the fact that visuospatial information, either in WM or LTM, was bound to the verbal digit information, facilitating performance.

Most studies of binding have focused on tasks that required participants to remember two features of a stimulus (e.g. shape and appearance): the bootstrapping task differs from these because it ostensibly requires participants to only remember a single (verbal) attribute, but allows in some conditions the opportunity of retaining information which can assist performance from the visuospatial modality. In this way the bootstrapping task is a method which can be used to observe implicit binding effects in a naturalistic setting. Similarly naturalistic studies have been conducted before entirely within the visuospatial domain: implicit extraction of redundancies within sets of stimuli facilitated performance in an ongoing visuospatial WM task (Brady, Konkle & Alvarez, 2009). A related finding is that made by Jiang, Olson and Chun (2000) indicating that spatial configuration information was an important component of performance on a colour memory task. Crucially, though, bootstrapping represents a linkage between theoretically separate visuospatial and verbal memory systems.

One key aspect of the EB proposal was its role in handling bindings between WM and LTM. Whilst our earlier study clearly indicated implicit binding of information in VWM and VSWM, it was suggestive but ultimately inconclusive as to the role of LTM. The study included a condition in which the digits were presented in a horizontal line from 0 – 9. There was no evidence of memory facilitation in this condition, an observation which was considered *likely* to implicate LTM, as it implied the presence of bindings with representations of numerical keypads. Had binding occurred between VWM and VSWM in the absence of LTM involvement, then the linear display should also have produced a bootstrapping effect. However, alternative explanations for these findings are possible. For example, the benefit of keypads over linear arrays could have been a consequence of the keypad incorporating a richer 2-dimensional spatial array, or the fact that tracing a spatial path through the 1 x 10 item linear display would involve far more path crossing and overwriting than in a 3 x 3 (+1) keypad display, leading to interference caused by trace overwriting.

The aim of the experiment reported in the current paper was to explore the nature of the ‘visuospatial bootstrapping’ effect, and in particular to establish more firmly whether it is attributable purely to visuospatial coding in temporary memory or whether LTM representations have a key role. Four display conditions were investigated: a single digit (SD) display and a typical keypad (TKP) display, replicating the conditions in which the original bootstrapping effect had been observed (Darling & Havelka, 2010). A third, novel static keypad (NSKP) display, employed a keypad shaped grid where digits were located randomly on the first trial, but then stayed in the same place throughout all trials in the condition. The purpose was to see if participants would be able to develop a LTM representation of the array over the course of several trials that would benefit their digit recall performance. Finally, as a strong test of the hypothesis that bootstrapping facilitation could occur between VWM and

VSWM with no influence of LTM, a novel changing keypad (NCKP) display was also used. This adopted a similar keypad, but digit-location mappings changed randomly with every presentation of an item, so there was no possibility of an existing LTM representation being present to facilitate recall performance. The four conditions are illustrated in Figure 1.

There is a long literature going back to Miller (1956) on interactions between LTM and WM focusing on the way that stored knowledge can integrate new material into ‘chunks’ which can facilitate storage. One of the purposes of developing the EB component was to specify a mechanism by which such ‘chunks’ might be formed or integrated (Baddeley et al., 2011). Should LTM be shown to have a critical involvement in the bootstrapping paradigm, though, this would go beyond traditional chunking by demonstrating the existence of processes that, on the basis of task demands, can implicitly bind together not just LTM and VSWM traces (e.g. Chase & Simon, 1973) or VWM traces (e.g. Miller, 1956), but instead LTM, VWM and VSWM traces in a complex multimodal representation that facilitates task performance.

Figure 1 around here

Method

Participants

Forty-eight participants (mean age = 23.5, SD = 8.13: 4 males) took part. All were staff or students attending Queen Margaret University (QMU).

Design

There were four display conditions for the digit span task: SD, TKP, NSKP and NCKP. All participants took part in all conditions in a fully counterbalanced order. The dependent variable was the proportion of trials in which all items were recalled correctly in their original order.

Materials and Procedure

Digit serial recall tasks.

The digit serial recall tasks involved immediate serial recall of a sequence of seven digits. In each sequence, the digits 0 - 9 were randomly sampled without replacement seven times. In any given condition, six practice trials were presented, followed by 24 experimental trials.

A standard laptop PC with a 1024 x 768 px display was used to present the stimuli. Each trial was initiated by the participant pressing a key, then a fixation cross was displayed for 500ms, followed by a blank screen interval of 250ms, and then digit sequence presentation. Digits were always presented in Arial Font, point size 18, and presented centrally within squares of side 60px.

In the TKP condition, the digits 0 - 9 were presented in the style of a traditional telephone keypad. The middle digit of the keypad was aligned centrally in the screen. There was a spacing of 12px between the outlines surrounding each digit. The sequence of digits was indicated in turn by highlighting the background of each digit, in green. Each digit was highlighted for 1000ms. Between items, the entire screen was cleared for 250ms. Once all seven digits had been indicated, the screen cleared for a 1000ms retention interval, and then the message 'RECALL' was presented in the middle of the screen. At this point participants attempted to verbally recall the digits of the sequence in the correct order.

The NSKP keyboard condition was identical to the TKP condition, with the digits appearing within the same grid as in the TKP condition, except that digits were arranged in one of four pseudorandom patterns (counterbalanced across participants) so that the digit-location mapping was unfamiliar. Digit positions stayed constant across all trials, so participants could learn the digit locations mappings as they progressed through the trials.

The NCKP display was also based on the TKP display, but digit locations changed on an item-by-item basis. Not only were digit positions randomised at the start of every trial, but also after every time a digit was presented within a trial. Hence there was no opportunity to establish long-term mappings of digits to locations.

In the SD condition, the 7 digits were presented one after the other in the middle of the screen, in a single square outline. Each digit was visible on a green background for 1000ms, and there was a 250ms interval between digits, during which the screen was blank.

Results

Five participants had to be excluded on the basis of being close to floor (total proportion correct across all 4 conditions $<.1$) or ceiling ($>.9$). Separate scores were derived for trials 1-12 and trials 13 – 24 of each block. Figure 2 shows memory performance across the display conditions. A 4 (condition) \times 2(trial subset) ANOVA demonstrated a significant main effect of display type ($F(3,126) = 11.71, p < .001, \eta_p^2=.22$). Direct comparisons showed that more sequences were correctly recalled in the TKP condition than in all other conditions (all $p_{\text{unadjusted}} < .0005$). There were no significant differences between performance in the SD, NSKP and NCKP conditions (all $p_{\text{unadjusted}} >.05$). There was also a main effect of trial subset, with performance on trials 1-12 being significantly worse than on trials 13-24 ($F(1,42) = 5.99, p = .019, \eta_p^2=.13$).

These two main effects need to be interpreted in the light of a significant interaction between trial subset and display condition ($F(3,126) = 3.07, p = .030, \eta_p^2=.07$). Planned comparisons between performance on Trials 1-12 and on Trials 13-24 indicated that participants recalled significantly more sequences correctly in the later trials than the earlier ones in the NSKP condition ($p = .001$) but not in the SD condition ($p = .307$), the TKP ($p = .451$) or the NCKP condition ($p = .368$).

Discussion

Replicating Darling and Havelka (2010), we found a significant and robust bootstrapping effect, with recall using the typical keypad display being significantly higher than all other conditions. In other words, immediate serial recall of digits is facilitated by incorporating additional spatial information available in the display. We estimate (based on the number of correctly recalled sequences) that additional visuospatial information can expand the effectiveness of digit sequence memory by around 9 %. There was no evidence that verbal memory was enhanced in either the NSKP or NCKP conditions, relative to the SD condition, but there was clear facilitation in the TKP condition. The only major difference between the TKP and the NSKP conditions is the fact that the standard typical keypad is a familiar stimulus that is likely to be represented in LTM. Thus these results demonstrate that the bootstrapping effect requires linkages of information between working memory and LTM. The NCKP condition facilitated a strong test of the hypothesis that an STM trace could support visuospatial bootstrapping in the absence of any LTM involvement, and the lack of any evidence of improved performance is a strong indicator that the bootstrapping effect requires at least some LTM involvement. Although this claim was asserted tentatively by Darling & Havelka (2010), on the basis of the comparison of linear and keypad stimuli, other explanations remained plausible. The current data permit the conclusion that bootstrapping represents the joining together of feature traces in VWM and VSWM utilizing LTM knowledge (of the typical keypad) to create ‘chunks’. In contrast with chunking reported previously in the literature (e.g. Chase & Simon, 1973; Miller, 1956) the LTM support is bound not just to a single type of STM material, but instead LTM, VWM and VSWM are all bound together.

The EB is held to be a single component that maintains bound representations from throughout memory, both LTM and STM, in the absence of attention or executive input

(Baddeley, et al., 2011). The majority of studies of binding that have been used to develop the EB model have been based on tasks that explicitly require binding between pairs of features held in WM. For example, Allen, Baddeley and Hitch (2006) required participants to remember the shape and colour of a set of objects, and Baddeley, et al. (2009) reported evidence of binding between verbal WM and LTM. However, the model suggests that integrated bindings of some complexity should be possible: in this context our identification of a binding between verbal, spatial and visual semantic (LTM) features is informative.

Participants' performance improved in the NSKP condition between the earlier and later trials. As can be inferred from Figure 2, performance on the initial 12 trials of the NSKP condition was *poorer* than in the baseline SD condition, whilst in later trials it was in line with performance in the SD and NCKP conditions. This pattern may suggest a conflict between an established digit-location mapping schema (the typical keypad) and a novel one, which was overcome as a new representation became established. It is also possible that experimentally induced interference caused by recent activation of the typical keypad representation might have impaired performance in the NSKP condition; however, a comparison between participants who took part in the TKP condition before the NSKP condition and those who did the TKP condition before the NSKP condition showed virtually identical levels of performance over the first 12 trials ($M_{TKP-NSKP} = .44$, $SD = .24$; $M_{NSKP-TKP} = .44$, $SD = .25$, $t = .005$, $p = .99$) suggesting that this was not the case. Consequently, it is likely that the source of this interference was the established presence in LTM of a keypad representation, rather than a consequence of its recent activation. A further alternative is that the regularity of digit positions in the NSKP condition was detected and that the pattern reflects an attentional cost of attempting to learn the novel verbal-spatial conjunctions.

The current results are potentially consistent also with the theoretical approach of Cowan (2005). The EB model itself is not inconsistent with the focus of attention approach,

so perhaps this is not surprising. One area where these approaches potentially differ is in terms of the role of attention in handling bindings across features: the EB is now (Baddeley et al., 2011) thought to be largely free from attentional involvement, whilst Cowan's approach implicates attention more. Given that our task probably invokes implicit binding (at least for TKP representations), it tends to favour the EB approach, but further research is certainly needed to strengthen this argument, as the possibility that attention is recruited in learning novel displays remains open.

The selective advantage for TKP displays over all other conditions in this study is the principal evidential basis for the strong claim we make that visuospatial bootstrapping tasks invoke LTM representations. However, given that such LTM representations do exist then the question is raised as to how they are established: it should be possible to experimentally create them *de novo*. It is clear that it would take more exposures than occurred in the current study, as performance in the later trials in the NSKP condition did not exceed the later trials in the SD condition. Future research should seek to characterize exactly what the parameters are by which these kinds of representations are learnt.

Overall, this experiment suggests that information held in separate subcomponents of working memory can be linked together to enhance performance on the verbal span task, and that information from representations in LTM can be used to achieve this. It is exactly this type of operation that is held to be a primary function of the EB (Baddeley, 2000, Baddeley et al., 2011). While binding in working memory (e.g. between shapes and colors, or letters and locations) has been reported previously, the observation of interactions between VWM, VSWM and LTM in this context is novel. These patterns emerge as a consequence of participants attempting an explicitly *verbal* memory task which had no explicit visuospatial or LTM content. The current study therefore defines the bootstrapping effect as requiring input from LTM representations in order to bind spatial and visual material.

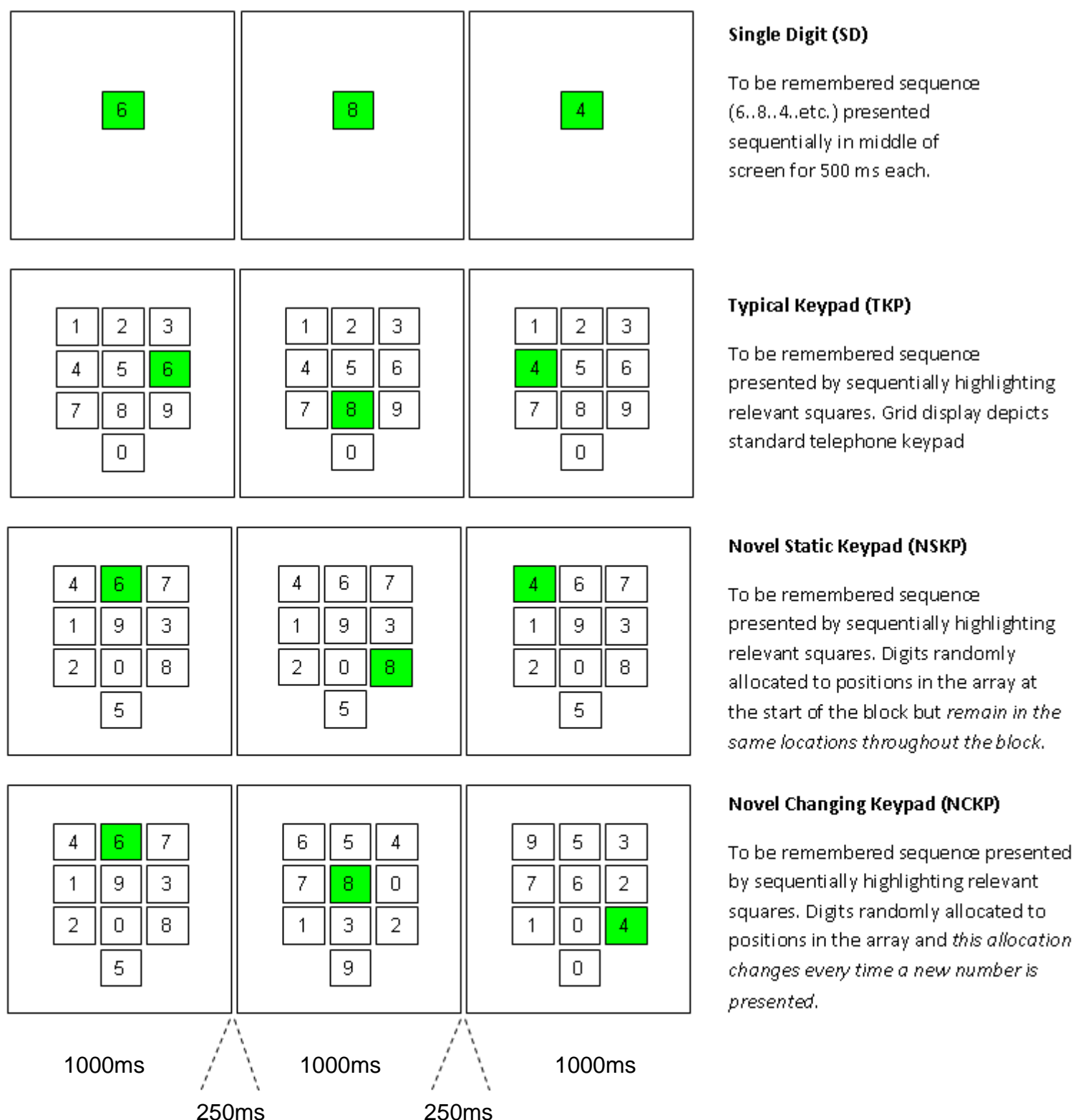


Figure 1: Figure depicting the difference between SD, TKP, NSKP and NCKP displays of the to-be-remembered digits (only first three digits of sequences are depicted in this figure).

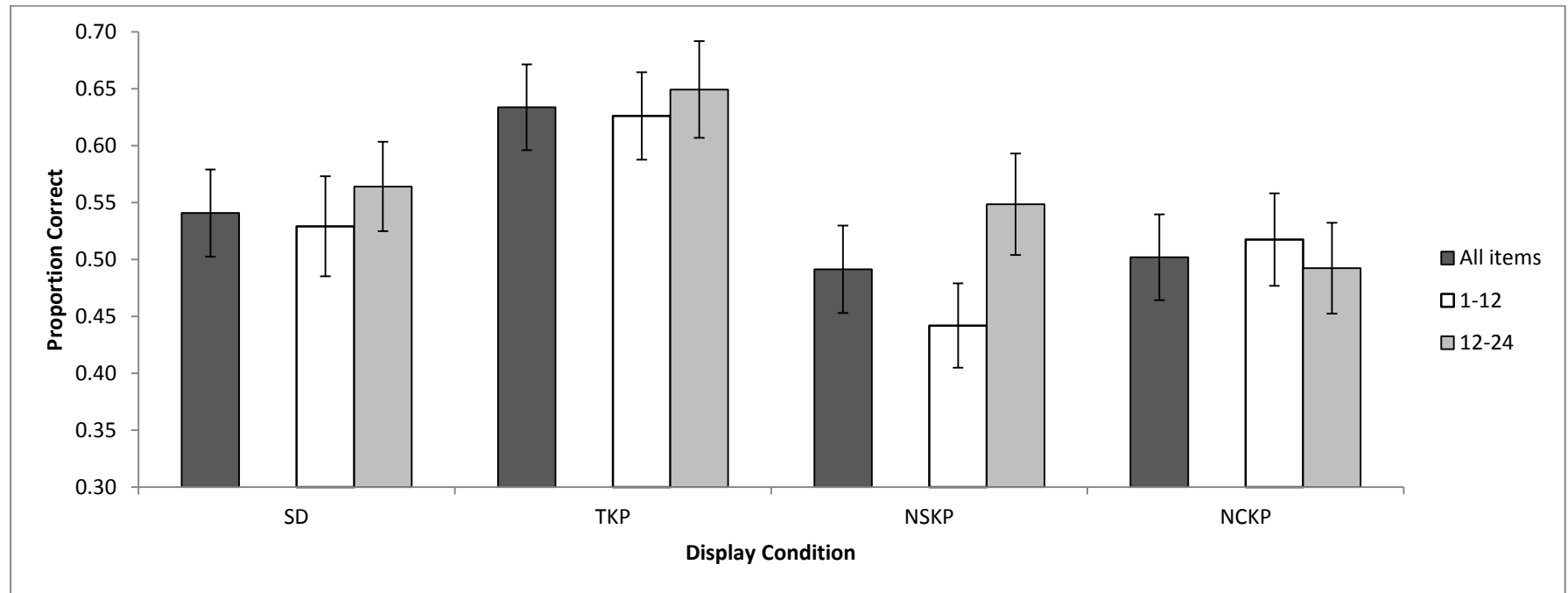


Figure 2. Memory performance (proportion of trials remembered correctly) across the 4 display conditions (shown by darkest bars). Breakdown of performance on trials 1 – 12 (empty bars) and 13 – 24 (light grey bars) is also shown. Error bars represent standard error of the mean.

References

- Allen, R. J., Baddeley, A. D., & Hitch, G. J. (2006). Is the binding of visual features in working memory resource-demanding? *Journal of Experimental Psychology: General*, 135, 298-313.
- Allen, R. J., Hitch, G. J., & Baddeley, A. D. (2009). Cross-modal binding and working memory. *Visual Cognition*, 17, 83-102.
- Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, 4, 417-423.
- Baddeley, A. D., Allen, R. J., & Hitch, G. J. (2011). Binding in visual working memory: The role of the episodic buffer. *Neuropsychologia*, 49, 1393-1400.
- Baddeley, A. D., Hitch, G. J., & Allen, R. J. (2009). Working memory and binding in sentence recall. *Journal of Memory and Language*, 61, 438-456.
- Baddeley, A., Lewis, V., & Vallar, G. (1984). Exploring the articulatory loop. *The Quarterly Journal of Experimental Psychology Section A*, 36, 233-252.
- Baddeley, A. D. & Hitch, G. J. (1974). Working Memory. In G.H.Bower (Ed.), *The Psychology of Learning and Motivation* (pp. 47-90). London: Academic Press.
- Bao, M., Li, Z. H., & Zhang, D. R. (2007). Binding facilitates attention switching within working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 959-969.
- Brady, T. F., Konkle, T., & Alvarez, G. A. (2009). Compression in visual short-term memory: using statistical regularities to form more efficient memory representations. *Journal of Experimental Psychology: General*. 138, 487-502

- Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin*, 104, 163-191.
- Cowan, N. (1998). Visual and auditory working memory. *Trends in Cognitive Sciences*, 2, 77-78.
- Cowan, N. (2005). *Working memory capacity*. Hove, East Sussex, UK: Psychology Press.
- Cowan, N., Sauls, J.S., & Morey, C.C. (2006). Development of working memory for verbal-spatial associations. *Journal of Memory and Language*, 55, 274-289.
- Chase, W. G. & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, 4, 55-81.
- Cocchini, G., Logie, R.H., Della Sala, S., MacPherson, S.E., & Baddeley, A.D. (2002). Concurrent performance of two memory tasks: evidence for domain specific working memory systems. *Memory and Cognition*, 30, 1086-1095.
- Darling, S. & Havelka, J. (2010). Visuospatial bootstrapping: Evidence for binding of verbal and spatial information in working memory. *Quarterly Journal of Experimental Psychology*, 63, 239-245.
- Darling, S., Della Sala, S., & Logie, R. H. (2009). Dissociation between appearance and location within visuo-spatial working memory. *The Quarterly Journal of Experimental Psychology*, 62, 417-425.
- Della Sala, S., Gray, C., Baddeley, A., & Wilson, B. (1997). *Visual patterns test: A test of short-term visual recall*. London: Thames Valley Test Company.

- Della Sala, S., Gray, C., Baddeley, A., Allamano, N., & Wilson, L. (1999). Pattern span: a tool for unwelding visuo-spatial memory. *Neuropsychologia*, 37, 1189-1199.
- Elsley, J. V. & Parmentier, F. B. R. (2009). Is verbal-spatial binding in working memory impaired by a concurrent memory load? *The Quarterly Journal of Experimental Psychology*, 62, 1696-1705.
- Jiang, Y., Olson, I.R. & Chun, M.M. (2000). Organization of visual short term memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 26, 683-702.
- Mate, J., Allen, R.J., & Baques, J. (in press). What you say matters: Exploring visual-verbal interactions in visual working memory. *The Quarterly Journal of Experimental Psychology*.
- Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 63, 81-97.
- Milner, B. (1971). Interhemispheric differences in the localization of psychological processes in man. *British Medical Bulletin*, 27, 272-277.
- Morey, C. C. & Cowan, N. (2004). When visual and verbal memories compete: Evidence of cross-domain limits in working memory. *Psychonomic Bulletin and Review*, 11, 296-301.
- Morey, C.C. & Cowan, N. (2005). When do visual and verbal memories conflict? The importance of working-memory load and retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 703-713.
- Pickering, S. J., Gathercole, S. E., Hall, M., & Lloyd, S. A. (2001). Development of memory for pattern and path: Further evidence for the fractionation of visuo-spatial memory.

The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 54, 397-420.

Prabhakaran V., Narayanan K., Zhao Z., & Gabrieli J.D. (2000). Integration of diverse information in working memory within the frontal lobe. *Nature Neuroscience*, 3, :85-90.

Quinn, J. G. & McConnell, J. (1996). Irrelevant pictures in visual working memory. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, 49, 200-215.

Smith, E. E., Jonides, J., & Koeppe, R. A. (1996). Dissociating Verbal and Spatial Working Memory Using PET. *Cerebral Cortex*, 6, 11-20.

Ueno, T., Mate, J., Allen, R.J., Hitch, G.J. & Baddeley, A.D. (2011). What goes through the gate? Exploring interference with visual feature binding. *Neuropsychologia*, 49, 1957-64.

Vandierendonck, A., Kemps, E., Fastame, M. C., & Szmalec, A. (2004). Working memory components of the Corsi blocks task. *British Journal of Psychology*, 95, 57-79.